Follicular lymphoma: is there an optimal way to define risk?

Carla Casulo
University of Rochester, Rochester, NY

Follicular lymphoma (FL) has a long natural history and typically indolent behavior. In the present era, there are a plethora of prognostic factors combining clinical, biological, and genetic data to determine patient prognosis and help develop treatment strategies over the course of a patient’s lifetime. The rapid pace of tumor-specific and clinical advances in FL has created a challenge in the prioritization and implementation of these factors into clinical practice. Developing a comprehensive understanding of existing prognostic markers in FL will help select optimal ways of utilization in the clinical setting and investigate opportunities to define and intervene upon risk at FL diagnosis and disease recurrence.

LEARNING OBJECTIVES
• Describe patient- and tumor-specific factors associated with risk of disease progression, histologic transformation, and premature death from FL
• Review established prognostic models in FL and their strengths and weaknesses when applied in clinical practice
• Discuss approaches to incorporate the clinical and mutational data into novel prognostic tools at different points in the patient’s lifetime

Introduction
In American and European adults, follicular lymphoma (FL) is the second most common non-Hodgkin lymphoma.1 The long natural history and typically indolent behavior of FL provide opportunities to identify prognostic information combining clinical, biological, and genetic data to determine patient prognosis and help define treatment strategies patients will need over their lifetime. However, the rapid pace of scientific discoveries in FL encompassing both tumor-specific and clinical advances has made it challenging to prioritize the most clinically relevant and essential prognostic parameters. Importantly, some of the prognostic information currently available often yields redundant or conflicting results, thereby limiting implementation into practice (Figure 1). A comprehensive overview of all prognostic markers to assess risk in FL is beyond the scope of this review, as it would be inclusive of genetic, genomic, clinical, and diagnostic imaging-based parameters.2-7 However, through a discussion of a patient case, this review examines some of the prognostic factors along the FL continuum, assesses optimal ways of utilization in clinical practice, and explores opportunities to define risk in the relapsed setting.

CLINICAL CASE
A 52-year-old man presented with painless cervical lymphadenopathy noticed while shaving. He has a history of hypertension and hyperlipidemia. He recently noted being significantly fatigued but denied weight loss or night sweats. Laboratory tests revealed hemoglobin of 10.2 g/dL (reference range, male: 13.7 - 17.5 g/dL), lactate dehydrogenase (LDH) of 270 U/dL (upper limit of normal, 225 U/dL), and an increased β2 microglobulin level (4.2 mg/L; upper limit of normal, 3.0 mg/L). No other abnormalities were noted. An excisional lymph node biopsy specimen showed grade 1 to 2 FL comprising small mature centrocytes in a well-preserved follicular pattern that stained positive for CD10, CD20, and BCL2. Flow cytometry detected a clonal population of B cells. Fluorescent in situ hybridization detected t(14;18). Bone marrow biopsy specimen revealed small paratrabeclular lymphoid aggregates encompassing 10% of the bone marrow space. Positron emission tomography (PET) showed hypermetabolic lymphadenopathy in the bilateral cervical chain, right axillae, retroperitoneum, and right inguinal area with standard uptake value of 7 to 11. Lymphadenopathy ranged from 3.2 to 4.7 cm.
Figure 1. Clinical prognostic tools and risk scores. ECOG, Eastern Cooperative Oncology Group; LN, lymph node; PS, performance status.
The patient initiated therapy with bendamustine and rituximab and received 6 cycles.

At present, there is no agreed-on, uniform definition of risk in FL. However, the survival of FL spanning over 2 decades and multiple available novel therapeutics with high response rates and long periods of remission provide a reasonable benchmark of expected patient outcomes at diagnosis. The concept of high-risk FL has emerged from data demonstrating increased probability of premature death for patients refractory to alkylator-based therapy or anti-CD20 monoclonal antibodies in the first-line setting, as well as patients with early disease recurrence following frontline therapy and early histologic transformation.8-10 For the purposes of this review, we consider patients at increased risk of these adverse outcomes as high risk.

An overview of the currently available clinical and tumor-based markers associated with high risk in FL reveals a multitude of variables (Table 1). Based on this patient’s male sex, presence of >4 nodal sites, high tumor burden, elevated LDH, advanced stage, and low hemoglobin, the patient has several factors associated with higher probability of poor outcome. Accordingly, the patient’s Follicular Lymphoma International Prognosis Index (FLIPI) score is 4, high risk, and based on this assessment, the patient’s overall survival (OS) is estimated to be 70% at 10 years for patients in the rituximab era.9,15 The FLIPI incorporates 5 clinical variables (nodal sites, elevated LDH, age >60 years, advanced stage, and hemoglobin <12 g/dL) into a multivariate model to predict OS. It was constructed to improve sensitivity compared with the International Prognostic Index for indolent lymphoma and based on retrospective data from patients diagnosed in the 1980s and 1990s treated heterogeneously at various centers.

Table 1. Variables associated with “high-risk FL” (increased risk of death, early recurrence, transformation)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Outcome</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical/patient based</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 3b vs grades 1-3a</td>
<td>Inferior OS</td>
<td>Wahlin et al,11 Mercadal et al12</td>
</tr>
<tr>
<td>Increased β2 microglobulin from normal</td>
<td>Inferior PFS and OS</td>
<td>Press et al,13 Bachy et al14</td>
</tr>
<tr>
<td>Large nodal mass (&gt;7 cm)</td>
<td>Inferior PFS</td>
<td>Solal-Céligny et al,15 Nooka et al,16 Buske et al17</td>
</tr>
<tr>
<td>Greater than 4 nodal sites</td>
<td>Inferior PFS and OS</td>
<td>Solal-Céligny et al15</td>
</tr>
<tr>
<td>Bone marrow involvement</td>
<td>Inferior PFS and OS</td>
<td>Bachy et al16</td>
</tr>
<tr>
<td>Low hemoglobin (under 12g/dL)</td>
<td>Inferior PFS and OS</td>
<td>Solal-Céligny et al,15 Nooka et al16</td>
</tr>
<tr>
<td>Advanced stage</td>
<td>Inferior PFS and OS</td>
<td>Brice et al,18 Buske et al17</td>
</tr>
<tr>
<td>High FL tumor burden based on GELF criteria</td>
<td>Inferior PFS and OS</td>
<td>Brice et al,18 Buske et al17</td>
</tr>
<tr>
<td>Low lymphocyte to monocyte ratio</td>
<td>Increased risk of transformation</td>
<td>Gao et al,19 Mohsen et al,20 Mozas et al21</td>
</tr>
<tr>
<td><strong>Imaging based</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High baseline SUV max of PET scan</td>
<td>Inferior OS and increased risk of early disease progression (mixed data)</td>
<td>Mir et al,22 Barrington and Meignan23</td>
</tr>
<tr>
<td>End of therapy PET positive (Deauville 4-5)</td>
<td>Inferior PFS and OS</td>
<td>Trotman et al7</td>
</tr>
<tr>
<td>High total metabolic tumor volume</td>
<td>Inferior PFS and OS (mixed data)</td>
<td>Barrington and Meignan,23 Skander 2019,50 Meignan et al6</td>
</tr>
<tr>
<td><strong>Tumor based</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23-gene GEP signature score</td>
<td>Inferior PFS</td>
<td>Huet et al5</td>
</tr>
<tr>
<td>Genomic alterations (gain of chromosome 2p; deletion 17p, subclonal TP53 mutations)</td>
<td>Inferior PFS</td>
<td>Qu et al26</td>
</tr>
<tr>
<td>Low intratumoral immune infiltration</td>
<td>Increased risk of early progression</td>
<td>Tobin et al25</td>
</tr>
<tr>
<td>High expression of genes from tumor associated macrophages</td>
<td>Inferior PFS and OS</td>
<td>Dave et al,24 Kridel et al27</td>
</tr>
<tr>
<td>Increased ctDNA at diagnosis and at completion of therapy</td>
<td>Inferior PFS and OS</td>
<td>Delfau-Larue et al,6 Zohren et al,28 Sarkozy et al29</td>
</tr>
<tr>
<td>Select mutations in several genes, including TP53, SOCS1, B2M, and MYD88</td>
<td>Enriched in tumors of early progressed patients</td>
<td>Kridel et al30</td>
</tr>
<tr>
<td><strong>Treatment based</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disease recurrence within 24 months of diagnosis or therapy</td>
<td>Inferior OS</td>
<td>Casulo et al,31 Jurinovic et al,32 Maurer et al33</td>
</tr>
<tr>
<td>Histologic transformation 18-24 months after diagnosis or therapy</td>
<td>Inferior OS</td>
<td>Link et al,4 Wagner-Johnson et al,36 Sarkozy et al35</td>
</tr>
<tr>
<td>Multiply relapsed disease</td>
<td>Inferior PFS</td>
<td>Batlevi et al,9 Link et al,24 Rivas-Delgado et al37</td>
</tr>
<tr>
<td>Refractory disease</td>
<td>Inferior PFS</td>
<td>Batlevi et al,9 Link et al,24 Rivas-Delgado et al37</td>
</tr>
</tbody>
</table>

GELF, Groupe d’Etude des Lymphomes Folliculaires; SUV, standard uptake value.
The FLIPI yields 3 risk classifications: low (0-1 factor), intermediate (2-3 factors), and high (4-5 factors) with a distribution of low, intermediate, and high risk in 36%, 37%, and 27% of patients, respectively. Despite its development in the prerituximab era, the FLIPI's prognostic legitimacy has been validated numerous times.16,17 This survival estimate may seem disproportionately low when considering recent studies evaluating prognosis and cause of death in FL. An analysis by Sarkozy et al11 noted that 10-year OS was approximately 75% in patients with newly diagnosed FL treated in the rituximab era, but a high-risk FLIPI score did increase the cumulative incidence of lymphoma-related mortality (incidence of lymphoma-related mortality at 10 years was 4.0% for low-risk FLIPI, 10.0% for intermediate-risk FLIPI, and 27% for high-risk FLIPI, P < .001). The patient's high-risk FLIPI score and high tumor burden therefore inform us of a likelihood of adverse outcome and the advantages of treatment initiation.

What are other clinical prognostic models available in the rituximab era?

As rituximab has drastically improved patient survival in the modern era, it is expected and observed that prognostic indices have less sensitivity and specificity for outcomes. The FLIPI2 was developed using 1093 patients with newly diagnosed FL in need of treatment, included nearly 70% of patients treated in the rituximab era, and was validated in an independent cohort of Italian patients.30 The FLIPI2 included 5 covariates based on their relative frequency and included elevated β2 microglobulin, lymph node diameter longer than 6 cm, bone marrow involvement, low hemoglobin, and age greater than 60 years. This model was designed to predict progression-free survival (PFS) and, if applied to this patient, would similarly yield a high-risk score of 3 and provide a 3-year PFS estimate of approximately 39% in the rituximab era. The validity of the FLIPI 2 was similarly confirmed externally in a cohort of 1135 patients from the PRIMA study who received chemotherapy with or without maintenance rituximab.

The Primary Rituximab and MAintenance (PRIMA) prognostic index (PRIMA-PI) estimated disparities in survival based on the number of factors present in patients treated in the PRIMA study.26 Using only 2 parameters (bone marrow involvement and increased β2 microglobulin), the primary end point was to predict PFS. In this case, the patient’s bone marrow involvement and increased β2 microglobulin would also result in a high-risk score but predict a different 5-year PFS of 69% and a 38% risk of experiencing early therapy failure (Table 2). Of importance, the PRIMA-PI was only predictive in patients treated with rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone or rituximab, cyclophosphamide, vincristine, and prednisone. Moreover, for patients with FL treated in the GALLLIUM study, bone marrow biopsy was not predictive of outcome.38 As with the FLIPI and FLIPI2, the PRIMA-PI model does not inform selection of therapy for this patient or modification of treatment based on his risk score.

### Table 2. Case patient’s risk assessment based on available clinical calculators/data

<table>
<thead>
<tr>
<th>Clinical Risk Calculator/Prognostic Data</th>
<th>Survival Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLIPI score high risk</td>
<td>10-year OS of 70%</td>
</tr>
<tr>
<td>FLIPI2 score high risk</td>
<td>3-year PFS of 39%</td>
</tr>
<tr>
<td>PRIMA-PI high risk</td>
<td>5-year PFS of 37%, 38% risk of having treatment failure within 24 months</td>
</tr>
<tr>
<td>FLEX score high risk</td>
<td>3-year PFS of 68%</td>
</tr>
<tr>
<td>Early relapse</td>
<td>5-year OS 50%</td>
</tr>
</tbody>
</table>

### Application of these models in the bendamustine/obinutuzumab era

The FLIPI, FLIPI-2, and PRIMA-PI were developed in patients treated with alkylators but did not include the more frequently used bendamustine. Although these indices have been validated retrospectively in small series, only the Follicular Lymphoma Evaluation Index (FLEX) model was predictive of outcomes in patients receiving bendamustine and the novel anti-CD20 antibody obinutuzumab.39 This model includes 9 variables: male sex, sum of products (SPD) in highest quartile, grade 3A, >2 extranodal sites, Eastern Cooperative Oncology Group performance status >1, hemoglobin <12 g/dL, β2 microglobulin higher than normal, peripheral blood natural killer cell count <100/µL, and increased LDH. A high score (3-9 factors) was predictive of poor PFS (3-year PFS of 86% for low risk vs 68% for high risk), and secondary end points also showed poor OS and increased risk of early progression. Given the lack of readily available variables tested in this model, as part of a sensitivity analysis, the authors also evaluated missing data. They found that the survival profile of patients with missing data appeared similar to that of complete cases, with a PFS hazard ratio for missing vs complete FLEX data of 0.89 (95% CI, 0.68-1.17; P = .40). In this patient’s case, his score would be 4 (high risk), based on sex, low hemoglobin, and increased LDH and β2 microglobulin. Other parameters included in the FLEX were not available for evaluation. This predicts a 3-year PFS of 86%. The clear disadvantage of this prognostic tool is the routine lack of several of the parameters calculated. In the FLEX manuscript, SPD was used as a surrogate for tumor burden. During model development, SPD had among the highest hazard ratios for PFS. As SPD requires radiologic calculation, bulky disease may be an alternative.

Although a Cox proportional hazards analysis of PFS showed that FLIPI and PRIMA-PI had a lower prognostic value across treatment regimens than the FLEX, it is not immediately apparent how this score will become routinely used in clinical practice, and further study is desirable in real-world settings and in the context of non-chemoimmunotherapeutic strategies such as lenalidomide and rituximab or rituximab monotherapy.

### Tumor-based, biologic, and genetic variables

FL cells reside within a rich tumor microenvironment (TME) enveloped by both nonmalignant cells and critical immune components that provide symbiotic influences on the FL tumor cell itself and back onto the TME. This contribution to patient outcome was first established by Dave et al,26 who demonstrated that high expression of genes from FL tumor-associated macrophages was indicative of poor outcomes. Yet, more contemporary studies show conflicting results when rituximab is incorporated into treatment regimens, as demonstrated by Kridel et al,27 who showed that high numbers of CD163+ macrophages were independent predictors of longer survival in patients receiving anthracycline-based regimens. In other work, the magnitude of intratumoral...
immune infiltration was found to be an important component of poor outcomes, in particular early treatment failure. A study from the Princess Alexandra Hospital evaluated whether there was an immune infiltration profile in FL associated with early disease progression.

Tobin et al performed targeted gene sequencing using NanoString technology from paraffin-embedded tissue and multispectral immunofluorescence on a tissue microarray that was applied to 2 groups: a discovery cohort of 132 patients from the Princess Alexandra Hospital with early and advanced stage FL who received either chemotherapy or observation and 2 independent validation cohorts of 198 patients with advanced stage disease treated with R-CHOP and R-CVP from the German Low Grade Lymphoma Study Group and the British Columbia Cancer Agency. They also performed T-cell repertoire analysis, flow cytometry, immunofluorescence, and next-generation sequencing. They showed that immune molecules from T cells, macrophages, or immune checkpoints were clustered into either high expression or low expression. Low levels of immune markers identified patients enriched for early progression, and PDL2 was the marker with highest accuracy to distinguish groups of low or high immune infiltration. This was validated in group of uniformly treated patients from the British Columbia Cancer Agency and the German Low Grade Lymphoma Study Group, which showed that tumors with a low immune infiltration had higher early disease-related progression of disease within 24 months of diagnosis/treatment of FL events. Immune infiltration by intratumoral T cells was quantified in conjunction with fluorodeoxyglucose-PET in a study by Nath et al. Lymph nodes from patients with high total metabolic tumor volume had an inverse association with numbers of intratumoral T cells but increased malignant B-cell infiltration. This suggests discrepant predictive utility of FDG-PET in patients with FL depending on the extent of T-cell depletion caused by the type of chemoinmunotherapy that is administered.

Gene expression profiling (GEP) studies by Huet et al defined a gene signature correlating with adverse PFS in FL. They defined a model based on the expression of 23 genes characteristic of B-cell centroblasts that was prognostic independent of the FLIPI score. In a multivariate analysis, those identified as high risk by the 23-gene predictor had a 5-year PFS of 26% compared with 73% in patients with a low-risk signature. This was confirmed in 3 independent validation cohorts. The performance of the 23-gene signature was evaluated in GALLIUM, in which no prognostic effect was observed for any of the gene signatures based on antibody treatment arm. However, chemotherapy choice did have a significant interaction between PFS and the 23-gene signature, in which patients with a high-risk score receiving CHOP/CVP had unfavorable outcomes compared with those treated with bendamustine, suggesting chemotherapy dependence of the high-risk signature score.

At present, the assessment of these immune markers and the capability to perform GEP studies are not standardized at diagnosis, and some techniques are not readily available and hence remain part of a research approach.

Combining clinical and genetic models

The combination of key epigenetic mutations with a high-risk FLIPI score and patient performance status was accomplished by Pastore et al to yield the m7-FLIPI. Although prior studies emphasized the relevance of single gene alterations, m7-FLIPI included a comprehensive collection of recurrent gene mutations as well as clinical risk factors in patients with FL. In the multivariate model, high-risk FLIPI score and Eastern Cooperative Oncology Group >1 carried significant weight, but EZH2 and ARID1A mutations identified a favorable risk population. The m7-FLIPI improved risk stratification by reclassifying patients previously classified as high-risk FLIPI into the low-risk m7-FLIPI group, which was a better reflection of their actual FL risk. However in a validation study of patients treated with bendamustine or obinutuzumab in the GALLIUM study, although the m7-FLIPI was prognostic in patients receiving CHOP/CVP and still outperformed the FLIPI, it lost its prognostic significance in patients receiving bendamustine or obinutuzumab. It is hypothesized that this interaction may be driven by EZH2 mutations, but this remains to be further validated. In nonchemotherapy-containing regimens, m7-FLIPI was also not prognostic of outcomes, raising the notion that chemotherapy dependence influences the high-risk signature score.

To evaluate the specific end point of POD24, Jurinovic et al reexamined the m7-FLIPI. Using the same patient cohorts and mutational data, a POD24-specific prognostic model was developed, described as the POD24-PI. It included FLIPI but was refined to include mutational status of only 3 genes (EP300, FOXO1, and EZH2). Compared with the m7-FLIPI, the new POD24-PI model had greater sensitivity to predict POD24 but was not superior to other approaches due to lower specificity. The Bio-clinical FLIPI (Bio-FLIPI) similarly integrated the FLIPI with genes implicated in the TME and identified several associated with early treatment failure. Nonetheless, only lack of intracellular CD4 expression was predictive of treatment failure. As in other examples, assessment of mutational status is not standardized and remains part of a research approach.

Dynamic risk assessment throughout the patient’s history

CLINICAL CASE (Continued)

The patient completed 6 cycles of treatment without complications. His PET response at the end of therapy was compatible with a complete metabolic response, an achievement associated with favorable outcomes. As part of an ongoing clinical trial, evaluation of minimal residual disease (MRD) was performed at diagnosis, end of therapy, and every 6 months for 2 years to assess correlation with disease progression.

Several studies have reported the association of MRD assessment as prognostic of outcome in FL. This can be achieved by assessment of circulating pretreatment cell-free DNA fragments from apoptotic tumor cells (circulating tumor DNA [ctDNA]) and by assessment of BCL2 immunoglobulin heavy chain levels quantified by polymerase chain reaction (PCR). Identification of the t(14;18) translocation can be qualitatively and quantitatively measured in the bone marrow and peripheral blood but is limited to patients harboring the t(14;18) and can also be found in the blood of healthy patients. Other technology using next-generation sequencing methods can detect a large spectrum of genetic alterations, such as the Cancer Personalized Profiling by Deep Sequencing method and quantitation of ctDNA encoding the clonal rearranged V(D) J Ig receptor gene sequence of FL cells.
A prospective German study analyzed BCL2 IGH levels by PCR in pre- and posttreatment peripheral blood in 173 patients with FL, finding an adverse impact of high pretreatment levels on subsequent PFS. As expected, high levels of MRD at the conclusion of therapy also had poor outcomes. The achievement of negative or low ctDNA or MRD similarly enhances prognostic information on patient outcome. Studies in other lymphomas have demonstrated that dynamic ctDNA levels during and after therapy not only can predict clinical outcomes but may also synergize with other prognostic factors such as total metabolic tumor volume. The best use of these molecular assessments of response will be during or following treatment to facilitate early intervention for recurrence, assess changes in mutational profile, and be used to guide next therapy.

**CLINICAL CASE (Continued)**

Approximately 1 year following therapy, the patient experienced disease relapse with high tumor burden, low-grade FL. The adverse prognostic impact of early disease recurrence (within 24 months of diagnosis) on survival has been well established. Few prognostic indices have been tested for use to predict outcomes at relapse outside of FLIPI, and this is an unmet need. He joined a randomized phase 2 clinical study randomizing patients to obinutuzumab with umbralisib, lenalidomide, or CHOP and was randomized to the obinutuzumab with umbralisib arm (NCT03269669). He experienced a partial metabolic response followed by rapid progression of disease. A repeat biopsy confirmed low-grade FL. He began therapy with axicabtagene ciloleucel on a clinical study (NCT03105336), now approved by the US Food and Drug Administration for relapsed FL in the United States (https://www.fda.gov/drugs/resources-information-approved-drugs/fda-grants-accelerated-approval-axicabtagene-ciloleucel-relapsed-or-refractory-follicular-lymphoma).

**Conclusions**

A multitude of prognostic variables are at our disposal to understand and predict survival for FL and to identify patients at high risk of poor outcome. However, several of these tools remain inaccessible in daily practice and have not been adequately tested to select therapy. Unfortunately, none of the risk models described here have been validated as a tool to select or adapt therapy for patients with FL. Although understanding prognosis is important, ultimately the goal of any of these tests should be to guide therapy. Studies of risk-adapted therapy based on MRD are ongoing in diffuse large B-cell lymphoma and beginning in FL (Luminari et al), but this approach remains investigational. Importantly, assessment of risk for the refractory patient or at relapse is an area of ongoing investigation and may contribute to a more precise ability to select the optimal treatment paradigm. The ideal way to merge existing clinical and biologic data in a sophisticated and accurate model for high-risk FL remains a work in progress.

**Conflict-of-interest disclosure**

Carla Casulo: Research funding from Gilead, Verastem, Genentech, and BMS.

**Off-label drug use**

Carla Casulo: nothing to disclose.

**Correspondence**

Carla Casulo, University of Rochester, 601 Elmwood Ave, Rochester, NY 14642; e-mail: carla_casulo@urmc.rochester.edu.

**References**


